Africa RISING

USAID's Sustainable Intensification Program in Africa
East and Southern Africa

Site selection in Malawi

March 22, 2013

Introduction

The Africa RISING program of the USAID-Feed the Future initiative proposes to initiate and test interventions to enable Sustainable Intensification (SI) of agriculture in three major regions of Africa by working in three “mega-sites” which exemplify the main climatic and human characteristics of these regions.

Malawi is one of the target countries in the East and Southern Africa mega-site. Dedza and Ntcheu districts (highlighted in Figure 1) have been selected as focused areas in the country. Figure 2 provides a further breakdown within the two districts to look at EPAs and the location of the four pairs of proposed sections (intervention and counterfactual) initially selected by MSU.

Figure 1. Dedza and Ntcheu districts in Malawi

Control sections (from which the control villages are selected) should be chosen within a pool of sections showing characteristics as much as possible identical to the action sections. Moreover, the control section must be physically or in some other way isolated from the action section. This procedure tries to avoid potential contamination effect[1] of the intervention sites into the control

[1] Contamination refers to the unintended effect of the intervention on the control sites, which then cannot be considered as valid counterfactuals. It is not to be confused with spill-over effect, which refers to the intended indirect consequence

sites. Contamination prevents a rigorous impact evaluation, possibly undermining the objective of the whole program, unlike spill-over effects that represent a desirable feature of the program, especially in Agriculture Research for Development (AR4D) projects.

Therefore, the distance between intervention and counterfactual sites becomes a crucial parameter to take into account in site selection. Table 1 displays the average distance between the pairs of intervention and counterfactual sections.

Figure 2. EPAs and locations of eight sections (4 intervention and 4 counterfactual sites) in Dedza and Ntcheu districts

of the intervention on neighboring villages, provided they have not been selected as valid counterfactual sites in the evaluation design phase.

Site	EPA	Section	Latitude	Longitude	Distance between Intervention and Counterfactual (km)
Dedza district	Linthipe	Mposa (Intervention)	14°12'21"S	34°05'58"E	
Dedza district	Linthipe	Ndikuwa (Counterfactual)	14°18'11"S	34°08'16"E	11.52
Dedza district	Golomoti	Golomoti (Intervention)	14°26'03"S	34°35'30"E	
Dedza district	Golomoti	Gosheni (Counterfactual)	14°20'45"S	34°40'01"E	12.7
Ntcheu district	Kandeu	Kampanje (Intervention)	14°37'45"S	34°35'51"E	
Ntcheu district	Kandeu	Sitolo (Counterfactual)	14°41'05"S	34°37'53"E	7.14
Ntcheu district	Nsipe	Mpamadzi (intervention)	14°55'47"S	34°44'47"E	
Ntcheu district	Nsipe	Nsipe (Counterfactual)	14°48'34"S	34°46'42"E	14.12

Table 1. Geographic characteristics of the selected sections in Malawi

From the table above, counterfactual and action sections are deemed to be too close within each pair to avoid contamination. The suggestion is to re-think the initial selection and choose the control sites in areas more distant from action sites within each of the strata, as they are constructed below.

Review of biophysical and socio-economic characteristics in Dedza and Ntcheu districts

In order to stratify and characterize the focused districts, a review of available spatial biophysical and socio-economic data layers is presented. The main purposes are to: 1. Understand the spatial pattern and homogeneity of each of the candidate data layers; 2. Choose the appropriate dataset for the stratification analysis.

The candidate layers are: population density, Agro-Ecological Zone (AEZ), precipitation, elevation, slope, farming system, market access, Length of Growth Period (LGP), and maize harvested area. The metadata of the individual datasets are listed in Table 2.

Datasets	Spatial resolution	Year	Source
Population density	1 sqkm	2000	CIESIN
Agro-Ecological Zones	~10sqkm		IIASA
Precipitation	50 sqkm	long term (> 50 years) average	CRU
	1 sqkm	long term (> 50 years) average	WorldClim
	100 sqkm	long term (> 50 years) average	NASA POWER
	50sqkm	long term (> 50 years) average	GPCC
	1sqkm	long term (1976-2008) average	interpolated from national weather station
Elevation	1 sqkm		USGS
Slope	1 sqkm		USGS
farming systems	shape file		John Dixon (2012 version)
Market access	1 sqkm	2000	HarvestChoice
Length of growth period	~10sqkm	long term (> 50 years) average	IIASA
Maize harvested area	~10sqkm	2000	HarvestChoice

Table 2. Characteristics of the candidate data layers

The variables are first mapped in order to visualize their spatial distribution, and then they are aggregated by classes.

1. Population density

Population density in Dedza and Ntcheu is generally higher than the average population density in East Africa. Most of the area shows population density higher than 100 persons per squared kilometer, and it is classified into 3 categories with the following cut-offs: less than 100, 100 -500, and greater than 500.

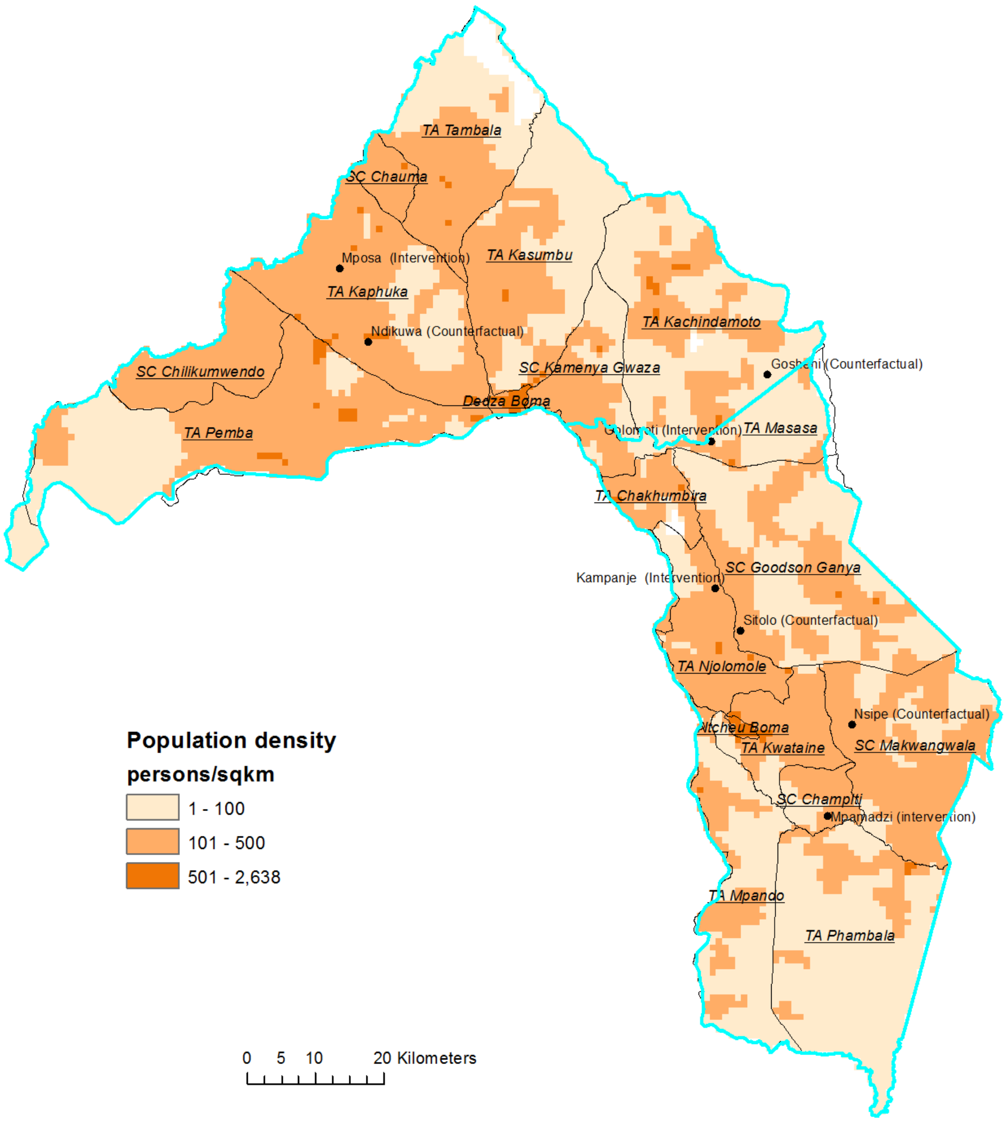

Figure 3. Population density

2. Elevation

There are many datasets available on elevation for Malawi: the USGS Hydro1k data layer has been chosen because most of the other data used in Africa Rising site selection analysis is at 1km resolution. In order to avoid arbitrary selection of cut-off values, the quintile of elevation distribution at 1km pixel level has been used.

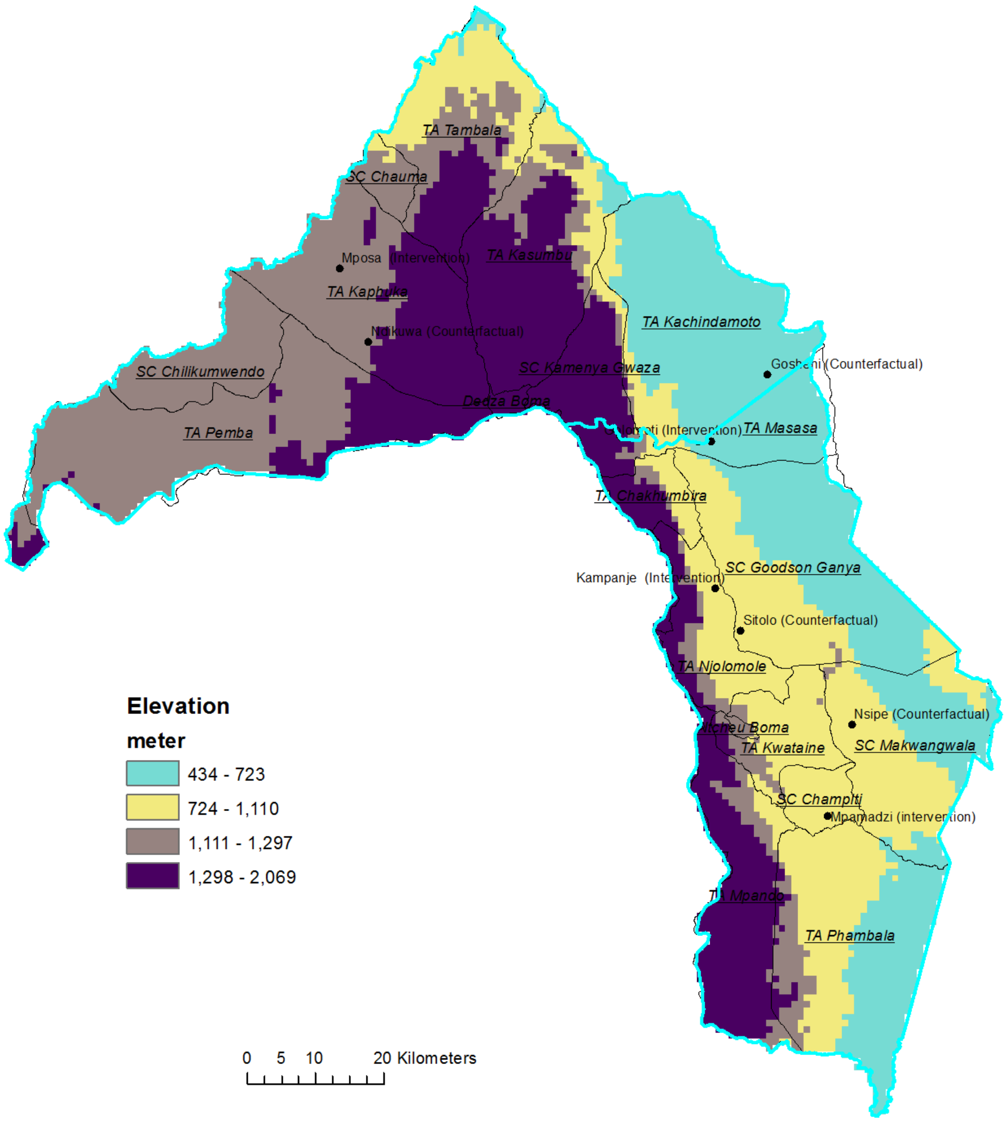

Figure 4. Elevation

3. Precipitation

Even though there are several publicly available precipitation data layers, most of them are more suitable to global studies than to country or sub-national analysis, being at a very coarse resolution. There are two methods to derive precipitation data point at the pixel level. One is from weather station records with spatial interpolation. The second method is from satellite observation.

The data from WorldClim has the highest spatial resolution, at 1km (Figure 5). Nevertheless, these data need to be used with caution, as their reliability has been questioned from various parts.

Other possible climatic data sources are CRU, NASA POWER, and GPCC, being all at half degree resolution. Their main drawback is the very coarse resolution, which makes them inadequate for the analysis on the focused districts.

Finally, the last option is relying on national weather data. The department of climate change and meteorological services at the Ministry of Natural Resources, Energy and Environment provide long term precipitation data from weather stations for the period 1961-2008 (see more at: http://www.metmalawi.com/weather/stations.php). IFPRI has access to these data through its office in Lilongwe, and the analysis below is based on the precipitation layer in *Figure 6* for the period 1976-2008 (Figure 5 is reported below just for comparison purposes).

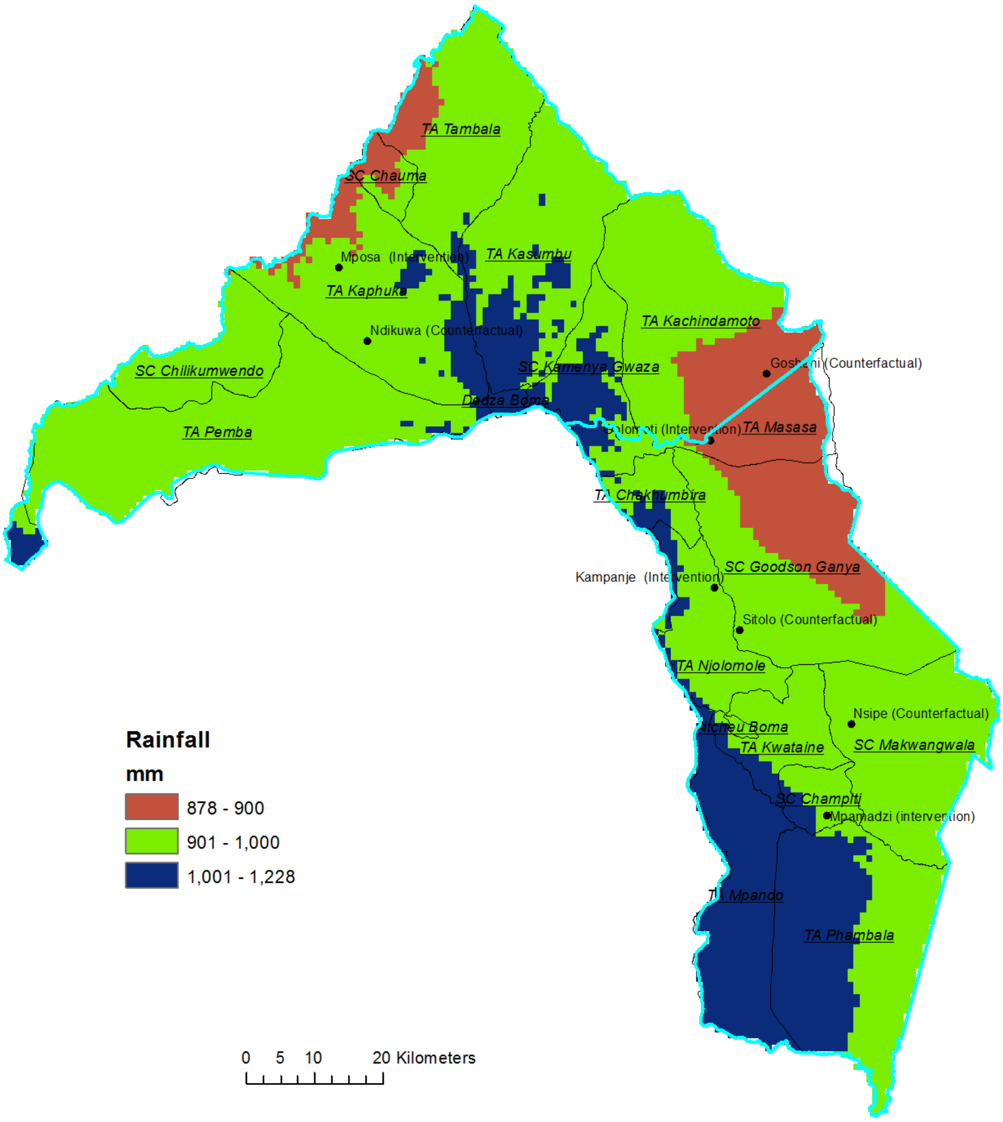

Figure 5. Long-term average precipitation (source: WorldClim)

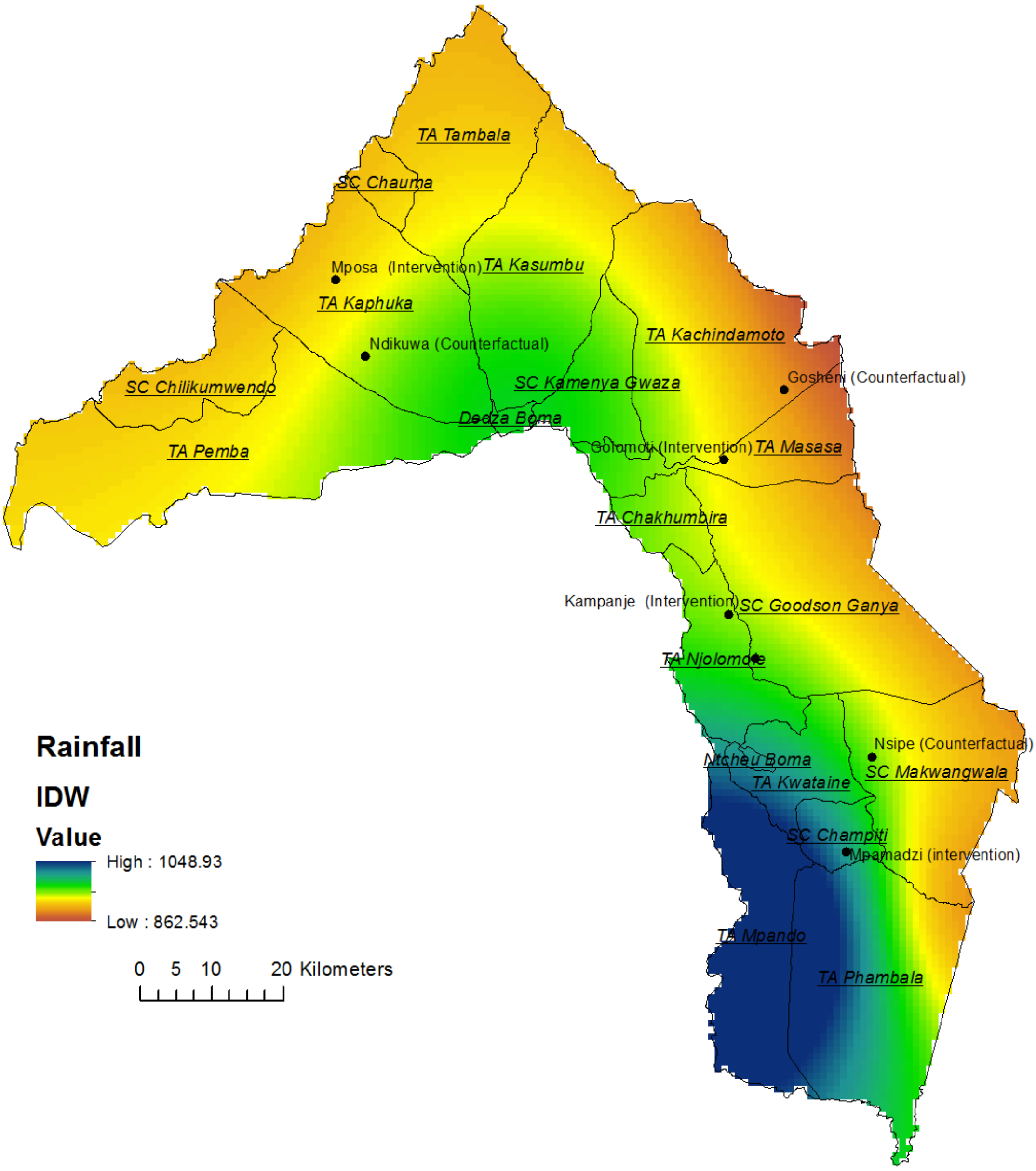

Figure 6. Long-term (1976-2008) average precipitation (source: Ministry of Natural Resources, Energy and Environment, Government of Malawi)

4. Market Access

Market access is largely used as an indicator of accessibility. The tercile classification (high, medium, and low) based on travel time in minutes to the nearest city with at least 50 thousand people is applied in this analysis.

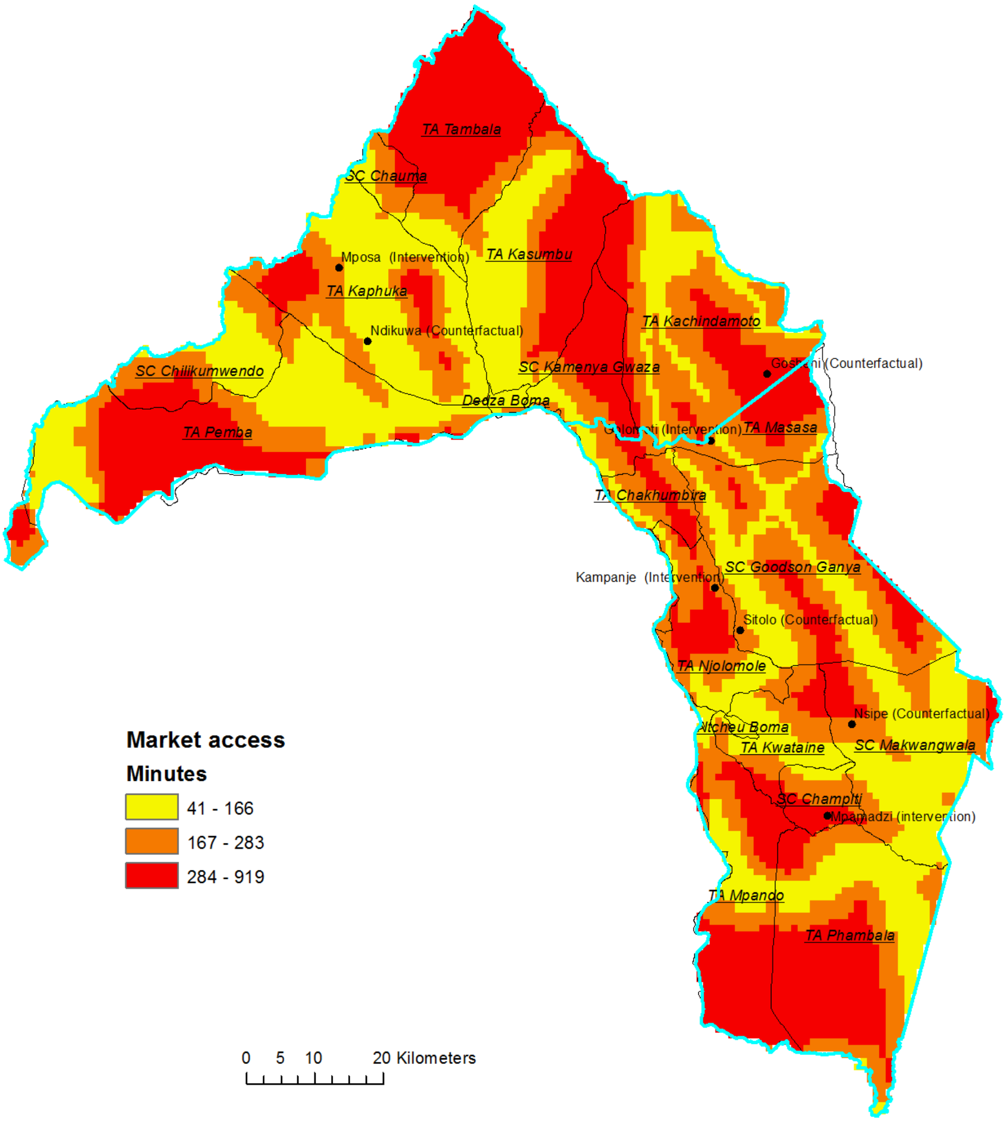

Figure 7. Market access

5. Slope

The slope is acquired from USGS Hydro 1k project. The quartile classification is used to classify the pixels.

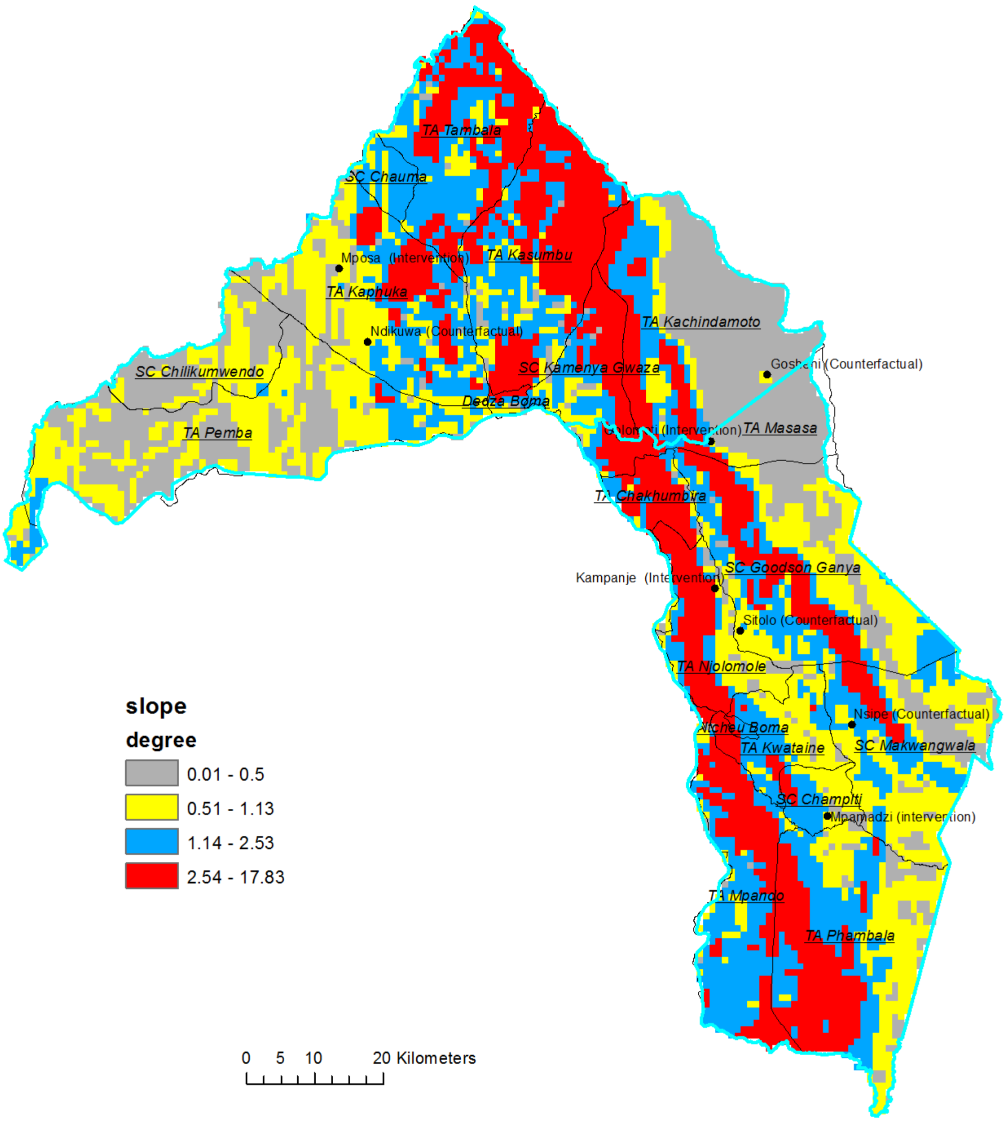

Figure 8. Slope

6. Maize Harvested area

Maize is the dominant crop grown in Malawi, and the maize harvested area is used in the analysis. The dataset is derived from the Spatial Production Allocation Model (SPAM) by HarvestChoice, where pixels are classified by quartile.

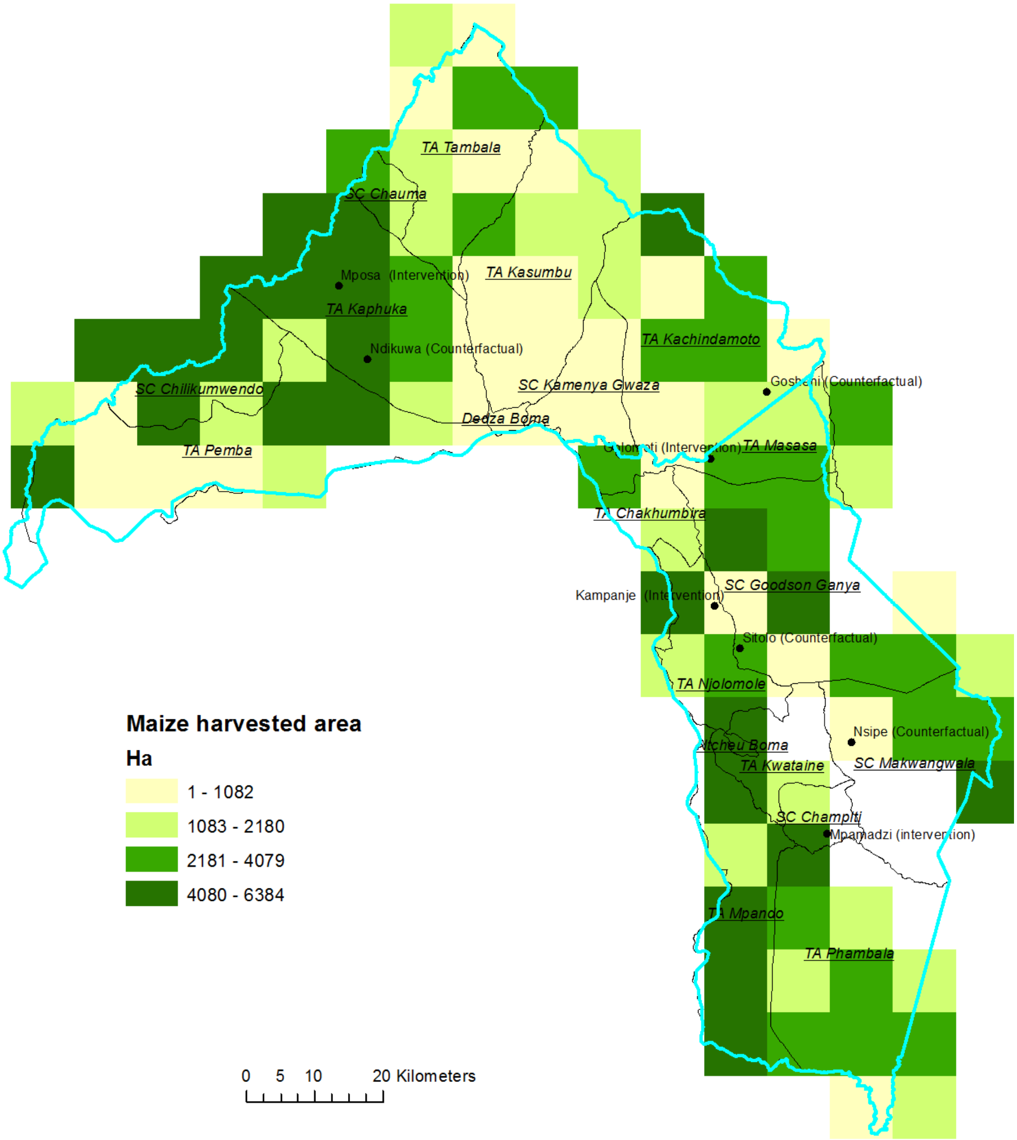

Figure 9. Maize harvested area

7. Length of growth period

The length of growth period, a good proxy of agriculture potential, measures how many continuous suitable days are available for the crop to grow based on soil water capacity holding, soil moisture, temperature, and elevation. Its quintile distribution is shown in Figure 10.

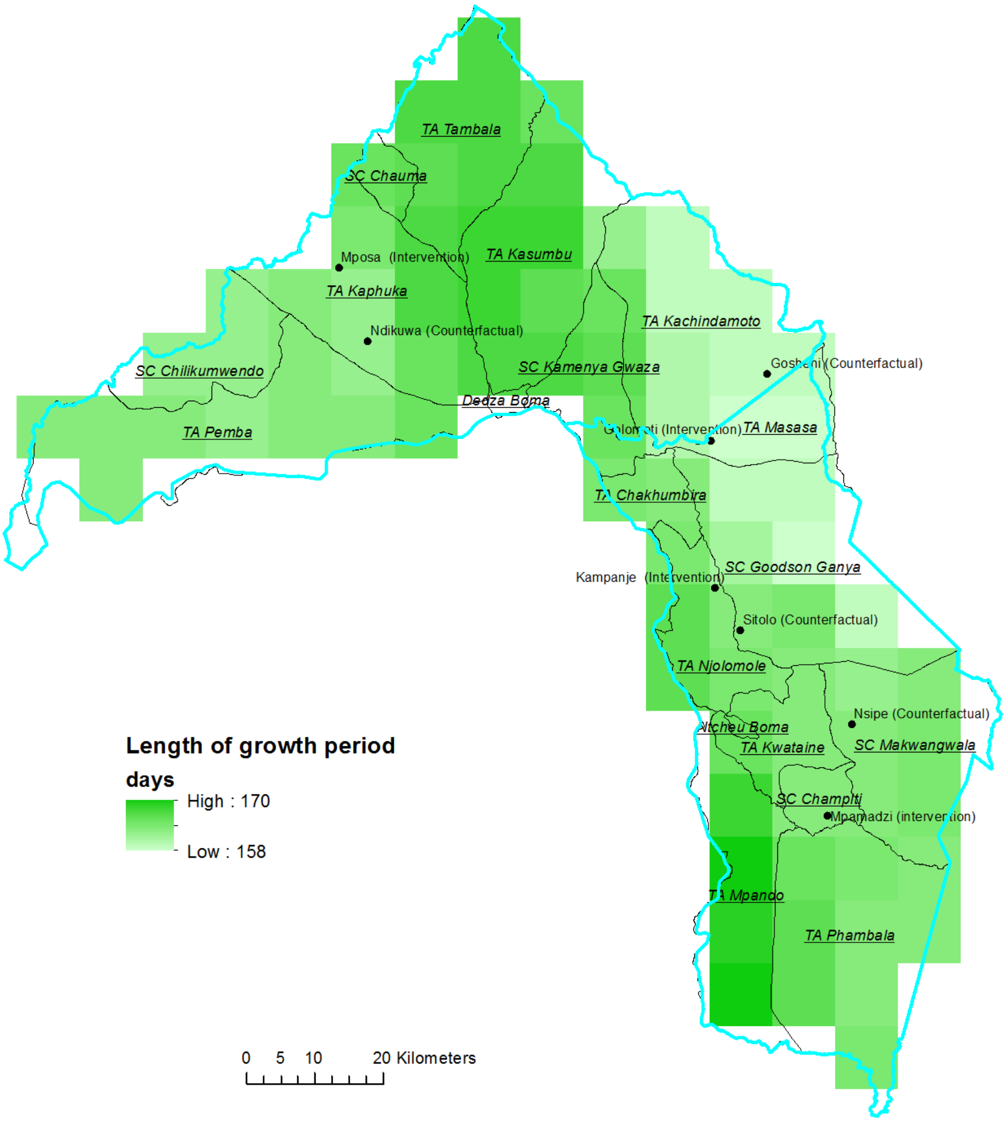

Figure 10. Length of growth period

8. Farming systems

The farming system map shows that the dominant system is maize mixed, with a small area based on artisanal fishing on the shores of Lake Malawi.

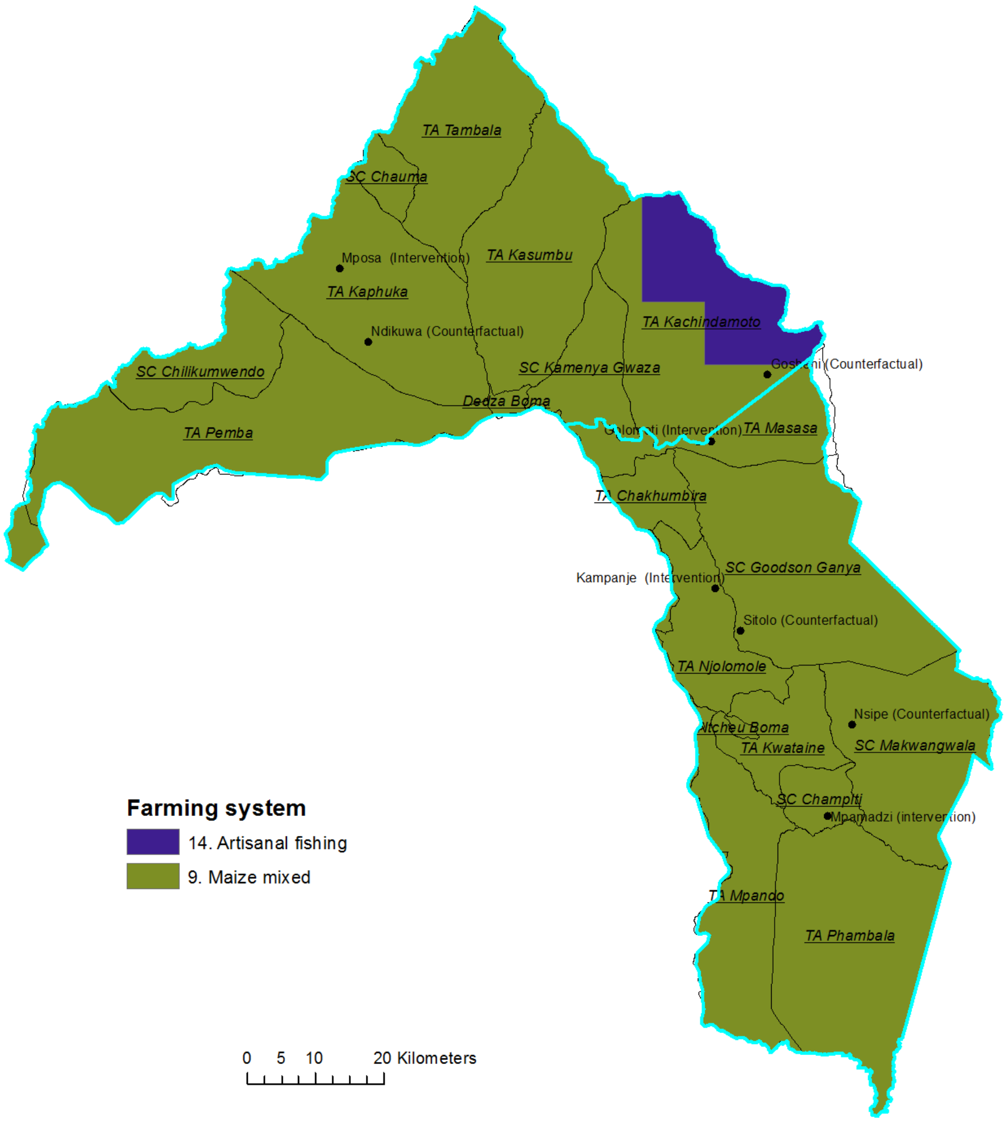

Figure 11. Farming systems

9. Agro-Ecological Zones

The AEZ layer shows two AEZ zones in Malawi: tropical cool semi-arid, and tropical warm semi-arid.

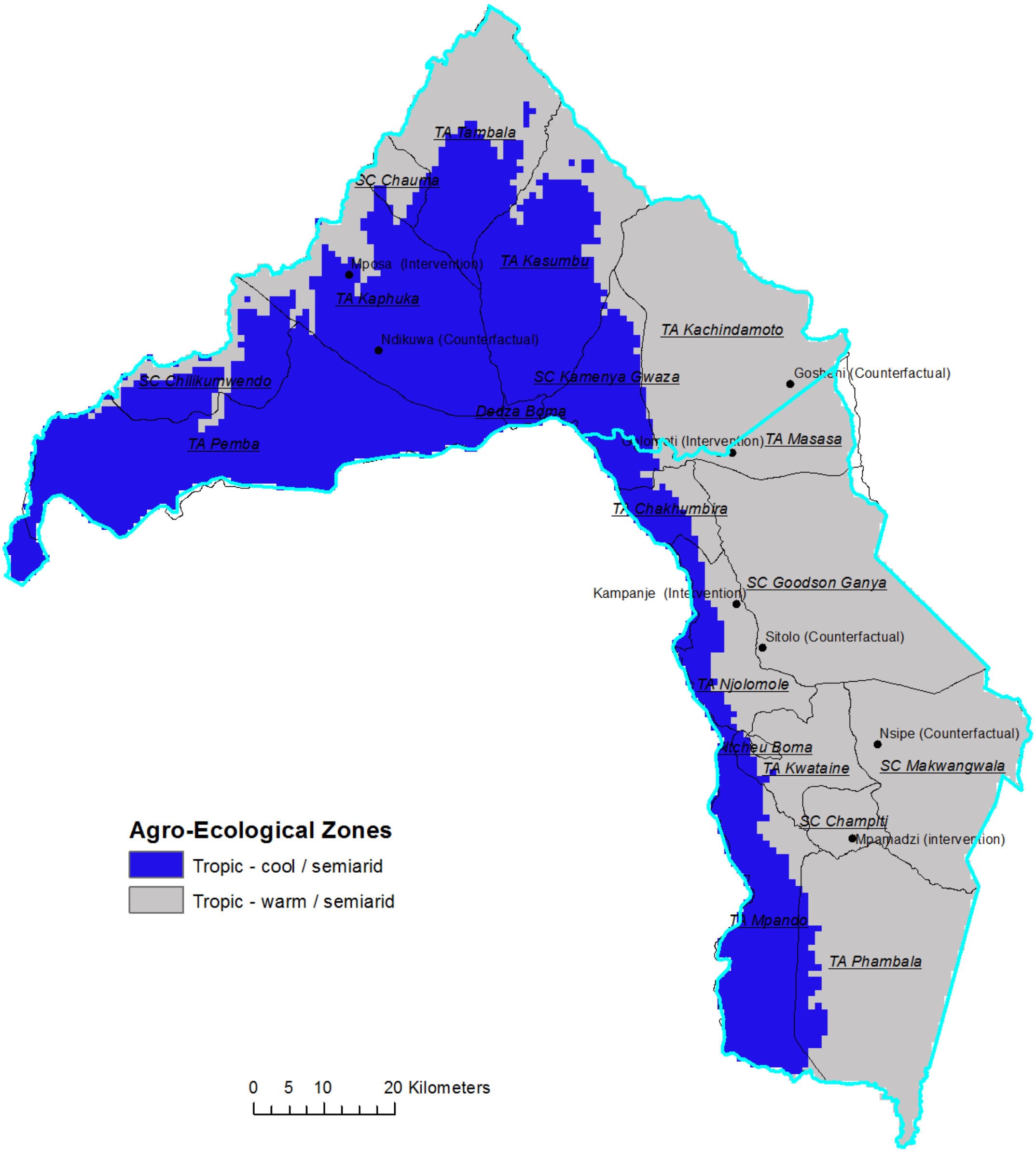

Figure 12. Agro-Ecological Zones

10. Surface temperature

The surface temperature layer is acquired from WorldClim. The long-term annual average temperature used in this report is shown in Figure 13. The spatial distribution of temperature is highly correlated with that of elevation, as it can be seen comparing *Figure 13* with *Figure 4*. The three categories based on temperature cutoffs (measured in degree Celsius, °C) of <20, 20-22, and >22 °C can be used to capture heat heterogeneity in the focused districts.

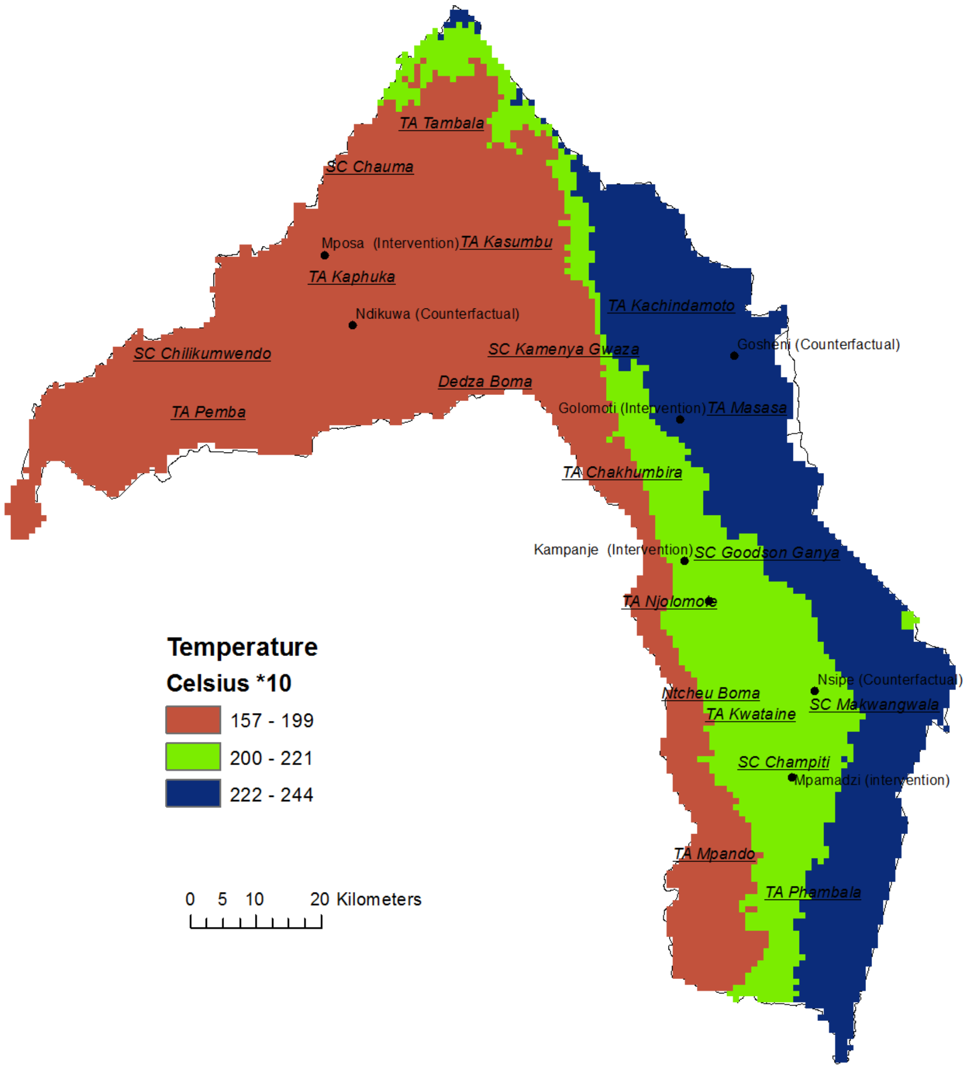

Figure 13. Surface temperature

Proposed stratification for site selection

After a review of the candidate variables, the data layers used to stratify the two districts are listed in Table 3.

class	pop density	rainfall	elevation	slope	market access	maize harvested area	Temperature
1	1-100	878-900	434-800	0.01-0.5	41-166	1-1082	15.7-19.9
2	101-500	901-1000	801-1000	0.51-1.13	167-283	1083-2180	20-22.1
3	500-2638	1000-1228	1000-2069	1.14-2.53	284-919	2181-4079	22.1-24.4
4				2.54-17.83		4080-6384	

Table 3. Proposed variables and their cut-offs

Elevation and rainfall, deemed as the two best proxies of agricultural potential for stratification purposes and representative of the variability of the main biophysical characteristics, are then used in the final classification process. The mean elevation at EPA level is calculated aggregating pixel values to EPA. Each EPA is then classified into High, Med, and Low elevation class based on information from Table 3 (Figure 14).

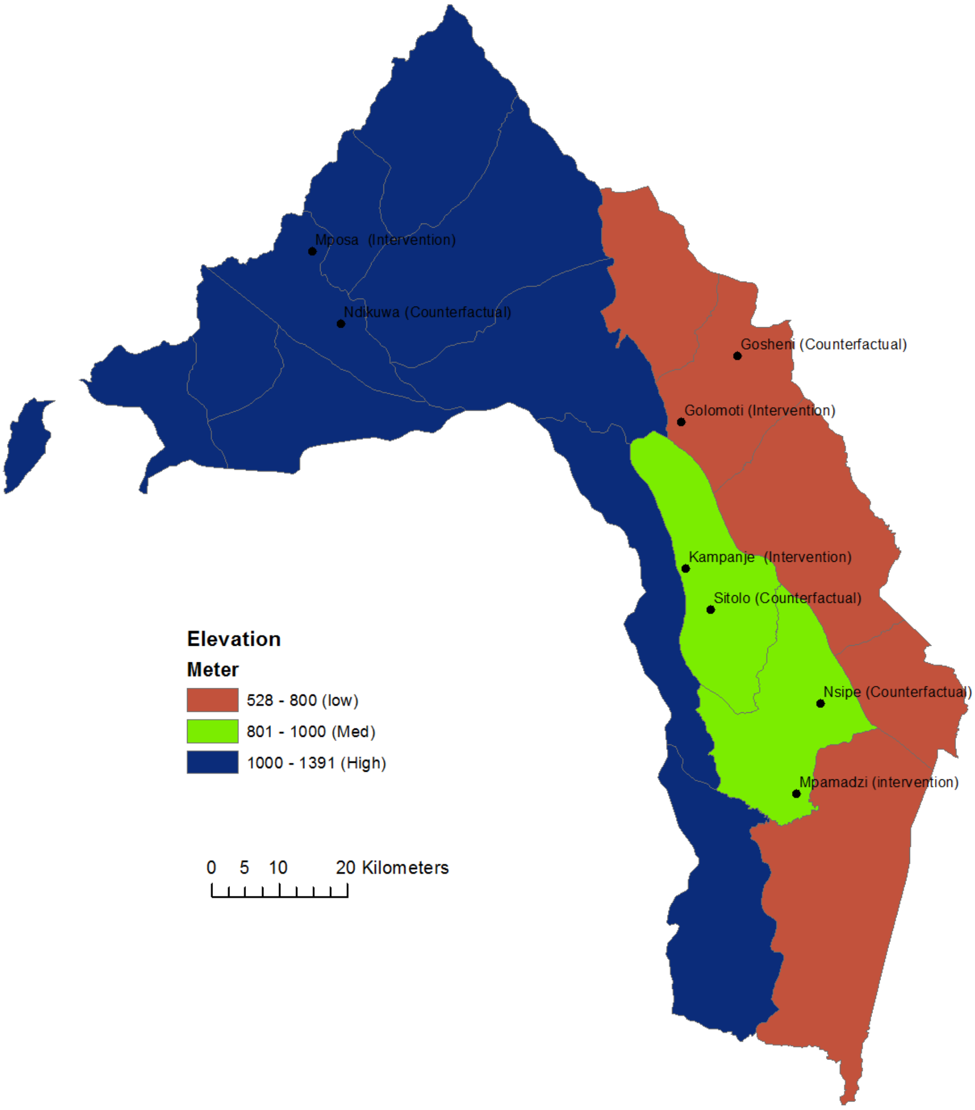

Figure 14. Elevation at EPA level

After expert consultations, it seemed apparent that none of the rainfall data available at HarvestChoice could adequately and reliably capture the spatial pattern on the ground. The team also estimated spatial interpolated values using data from the national weather stations, although the exercise was reckoned to be inadequate due to the paucity of observation points. Besides data quality issues, the evidence showed that using rainfall variables alone may not satisfactorily capture the water impact on crop growth, since water use efficiency was not known. It was then decided to combine rainfall and temperature data, and adjust the rainfall data based on temperature information.

The steps involved in the adjusted rainfall layer are described below, and results are displayed in Figure 15:

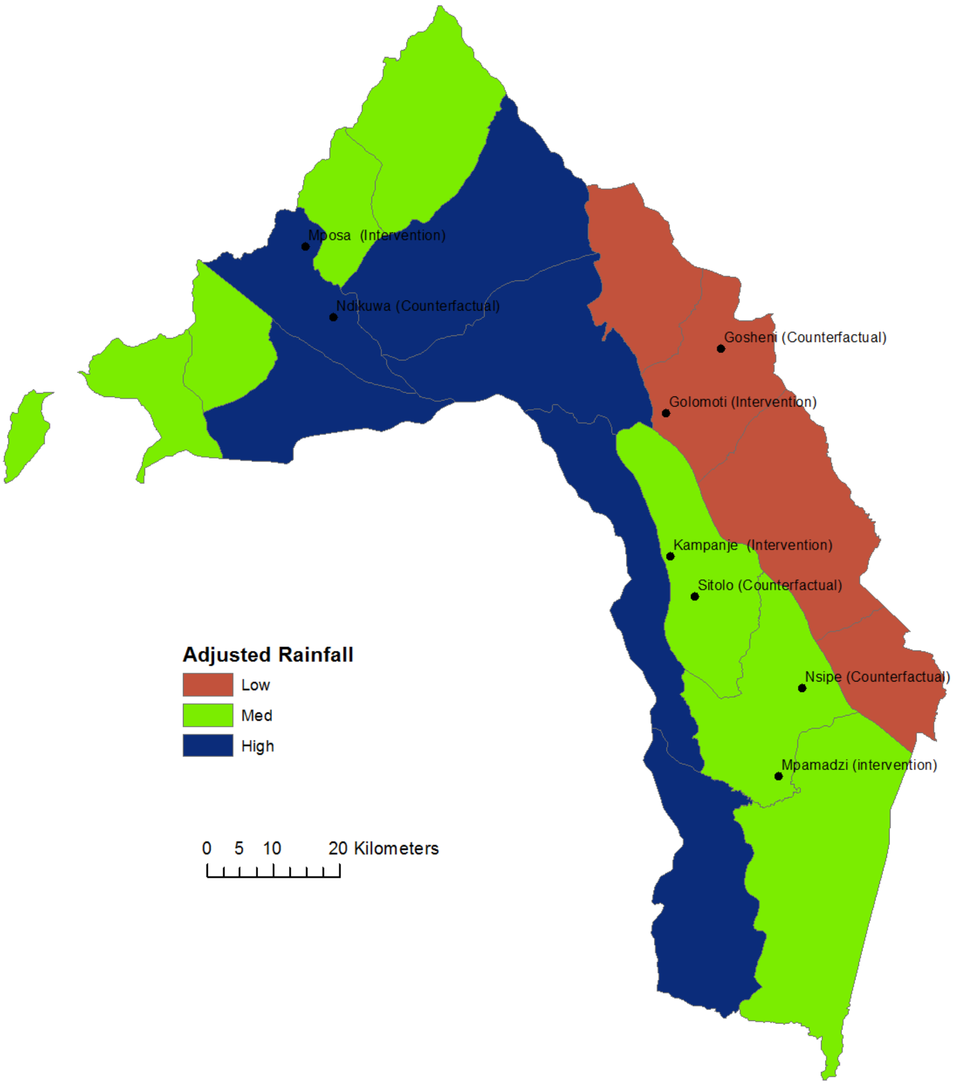

Figure 15. Adjusted rainfall at EPA level

1. Classify rainfall and temperature separately and assign value to each class as below:

	Rainfall class	Temperature class
High	3	1
Med	2	2
Low	1	3

2. Combine rainfall and temperature with equal weight at pixel level, so that the value of combined class is equal to [value of rainfall class + value of temperature class] / 2. As examples, for a pixel with high rainfall (3) and low temperature (3), the adjusted rainfall is equal to 3 ([3+3]/2; while for a pixel in med rainfall (2) and low temperature (3), the adjusted rainfall is equal to 2.5 ([2+3]/2).[2]
3. Use the combination as per point 2. to reclassify rainfall into High/Med/Low category at pixel level, with High class for value above 2.5; Med class for value between 1.5 and 2.5; and Low class for value lower than 1.5.
4. Assign each EPA the category High/Med/Low according the largest area (number) of pixels in that category by EPA.

The final classification is obtained by combining elevation unfeasible and adjusted rainfall class at the EPA level (Figure 16).

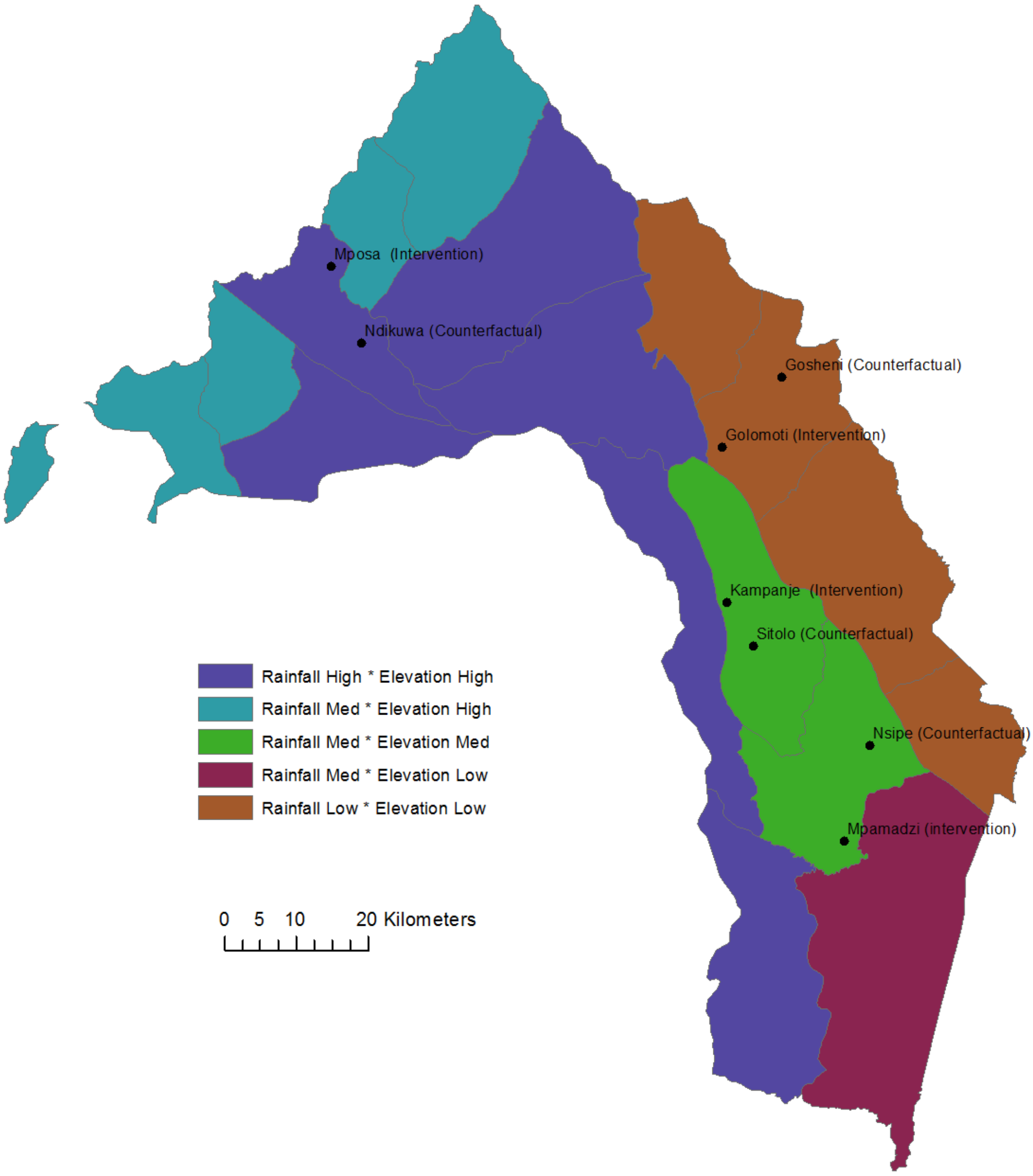

Figure 16. Rainfall and elevation classes at EPA level

[2] Note that some combinations, although theoretically possible, are extremely unlikely to occur on the ground. For example, there are very few pixels, classified in the med class (2), with high rainfall (3) and high temperature (1) or with low rainfall (1) and low temperature (3), due to the spatial distribution of rainfall and temperature in the two districts, shown in Figure 5 and Figure 13, respectively.

Classification results are also provided in tabular format in Table 4.

EPAs	Zones	District	Rainfall class	Elevation class	Rainfall * Elevation class
BEMBEKE	DEDZA HILLS	Dedza	Rainfall High	Elevation High	Rainfall High * Elevation High
CHAFUMBA	THIWI LIFIDZI	Necheu	Rainfall Med	Elevation High	Rainfall Med * Elevation High
KANYAMA	DEDZA HILLS	Dedza	Rainfall High	Elevation High	Rainfall High * Elevation High
KABWAZI	THIWI LIFIDZI	Dedza	Rainfall Med	Elevation High	Rainfall Med * Elevation High
LOBI	THIWI LIFIDZI	Dedza	Rainfall High	Elevation High	Rainfall High * Elevation High
LINTHIPE	THIWI LIFIDZI	Dedza	Rainfall High	Elevation High	Rainfall High * Elevation High
KAPHUKA	DEDZA HILLS	Dedza	Rainfall Med	Elevation High	Rainfall Med * Elevation High
MAYANI	DEDZA HILLS	Dedza	Rainfall Med	Elevation High	Rainfall Med * Elevation High
TSANGANO	NTCHEU	Necheu	Rainfall High	Elevation High	Rainfall High * Elevation High
NJOLOMOLE	NTCHEU	Necheu	Rainfall High	Elevation High	Rainfall High * Elevation High
MANJAWIRA	NTCHEU	Necheu	Rainfall Med	Elevation Low	Rainfall Med * Elevation Low
BILIRA	BWANJE VALLEY	Necheu	Rainfall Low	Elevation Low	Rainfall Low * Elevation Low
SHARPEVALE	BWANJE VALLEY	Necheu	Rainfall Low	Elevation Low	Rainfall Low * Elevation Low
MTAKATAKA	BWANJE VALLEY	Dedza	Rainfall Low	Elevation Low	Rainfall Low * Elevation Low
GOLOMOTI	BWANJE VALLEY	Necheu	Rainfall Low	Elevation Low	Rainfall Low * Elevation Low
KANDEU	NTCHEU	Necheu	Rainfall Med	Elevation Med	Rainfall Med * Elevation Med
NSIPE	NTCHEU	Necheu	Rainfall Med	Elevation Med	Rainfall Med * Elevation Med

Table 4. Classification of EPAs

Conclusions

Given the close proximity between action and counterfactual sections within each pair of sites identified by MSU, and the new stratification proposed in this document, a re-selection of sites in Ntcheu and Dedza districts in Malawi is advised.
In particular, Mposa and Ndikuwa in Lithipe EPA located in the high rainfall/high elevation domain should be considered both treated, and they should be paired up with other two (or more) sections within the same domain, either in the EPAs of Kanyama, Bembeke or in the south-western EPAs. In the same fashion, Gosheni and Golomoti in Golomoti EPA should be both considered as intervention sections, and they should be paired up with other sections in the EPAs of Sharpevale or Bilira, being in the low rainfall/low elevation class. Finally, Kampanje and Sitolo in Kandeu EPA should be targeted sections, and Nsipe, Mpamadzi in Nsipe EPA should be considered as counterfactual sections (or the other way around, if this is deemed easier based on practical considerations), being all within the medium rainfall/medium elevation domain. Moreover, sections in the EPA of Mayani and Kaphuka should be selected and paired up with corresponding sections in Chafumba and Kawazi, being within the medium rainfall/high elevation domain.
This selection strategy would target four strata/domains, leaving out only the low rainfall/med elevation domain -not of interest to Africa RISING-, allowing an adequate coverage of the spectrum of biophysical conditions prevailing in the two districts and a broad assessment of the interventions in areas with different agricultural potential.